On The Duality Between State-Dependent Channels and Wiretap Channels

David Kibloff, Samir M. Perlaza, Guillaume Villemaud, Leonardo S. Cardoso

ON THE DUALITY BETWEEN STATE-DEPENDENT CHANNELS AND WIRETAP CHANNELS

David Kibloff, Samir M. Perlaza, Guillaume Villemaud and Leonardo S. Cardoso

Université de Lyon, Inria, INSA Lyon, CITI, F-69621 Villeurbanne, France
E-mail: david.kibloff@inria.fr

ABSTRACT

In this paper, a duality between wiretap and state-dependent channels with non-causal channel state information at the transmitter is established. First, a common achievable scheme is described for a certain class of state-dependent and wiretap channels. Further, state-dependent and wiretap channels for which this scheme is capacity (resp. secrecy capacity) achieving are identified. These channels are said to be dual. This duality is used to establish the secrecy capacity of certain state-dependent wiretap channels with non-causal channel state information at the transmitter. Interestingly, combatting the eavesdropper or combatting the lack of state information at the receiver turn out to be two non-concurrent tasks.

1. INTRODUCTION

The wiretap channel (WTC) model was introduced by Wyner in [1]. Yet simple, the WTC captures the essential tradeoff between the point-to-point information rate between a legitimate transmitter-receiver pair and the normalized equivocation at a malicious receiver eavesdropping upon the former. Provided that the eavesdropper is physically degraded with respect to the legitimate receiver, Wyner reported the existence of at least one coding scheme able to simultaneously satisfy two tasks. First, reliably transmitting information between the transmitter and the legitimate receiver; and second, guaranteeing that the normalized equivocation at the eavesdropper is arbitrarily close to the normalized entropy of the message index. That is, guaranteeing that the information leakage is arbitrarily close to zero. These two tasks can be simultaneously fulfilled at least in the asymptotic block-length regime using a random binning technique that takes advantage of the degraded signal observation at the eavesdropper with respect to the legitimate receiver. Wyner fully characterized the rate-equivocation region of the WTC. Csiszár and Körner then fully described the rate-equivocation rate region of the broadcast channel with confidential messages in [2], when the eavesdropper is not necessarily physically degraded with respect to the legitimate receiver. Liang and Poor obtained similar results for the multiple-access channel with confidential messages [3].

The state-dependent channel (SDC), introduced by Shannon in [4], is a time-varying point-to-point channel. Therefore, several scenarios can be foreseen with respect to the channel state information (CSI): CSI not available at the transmitter and the receiver (no CSI); CSI available at both the transmitter and the receiver (full CSI); CSI available either at the transmitter or the receiver (CSI-T and CSI-R, respectively). In any of these four cases, the channel capacity might be different, which highlights the critical impact of CSI. Moreover, in each case, CSI can be causally or non-causally available. In [4], Shannon considered three cases: no CSI, causal CSI-T, and full CSI. Non-causal CSI was introduced later by Kuznetsov and Tsybakov [5] and generalized to discrete memoryless channels by Gelfand and Pinsker [6]. Interestingly, when CSI-T is available non-causally, the coding scheme used to achieve the channel capacity also consists of a binning technique.

In this paper, the links between these two models are investigated. In particular, it is shown that for a certain class of SDC and WTC pairs, there exists a common coding scheme for which any achievable rate for the SDC is achievable for the WTC and *vice-versa*. In this case the SDC's capacity equals the secrecy capacity of the WTC and the two channels are said to be dual.

The remainder of this paper unfolds as follows. Section 2 and Section 3 detail respectively the SDC and the WTC models. Section 4 describes the main results and Section 5 presents the corresponding proofs. Section 6 concludes this work.

Notation: Throughout this paper, random variables are denoted by an uppercase letter, e.g. X. The realizations of a random variable are denoted by a lowercase letter, e.g. x. The set of events is denoted by a calligraphic uppercase letter, e.g. $\mathcal{X}$. In general, the probability distribution of the random variable X is denoted by P_X, and belongs to the set of possible distributions P_X, denoted by $\triangle(\mathcal{X})$. Whenever a second random variable Y is involved, P_{XY} and $P_{Y|X}$ denote respectively the joint probability distribution of (X, Y) and the conditional probability distribution of Y given X. In case the random variable is an n-length vector, it is denoted by a boldface uppercase letter, e.g. $\boldsymbol{X} = (X_1, X_2, ..., X_n)$, and the realizations are denoted by a boldface lowercase letter, e.g. $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in \mathcal{X}^n$. The set of n-length typical sequences $\boldsymbol{X}$ is denoted by $\mathcal{T}_\epsilon^{(n)}(X)$. The expected value of the random variable X according to the distribution P_X is denoted by $\mathbb{E}_X[\cdot]$. Given a discrete memoryless SDC denoted by $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$, its corresponding capacity is denoted by $C_{ab}(K_S)$, where $(a, b) \in \{0, 1\}^2$, and $a = 1$ (resp. $b = 1$) indicates the availability of non-causal CSI at the transmitter (resp. the receiver). Alternatively, $a = 0$ (resp. $b = 0$) indicates the absence of CSI at the transmitter (resp. the receiver). Given a discrete memoryless WTC denoted by $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_{Y\,Z|X})$, its secrecy capacity is denoted by $C_s(K_W)$. Given a discrete memoryless state-dependent degraded WTC, denoted by $K_{SW} = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_S, P_{Y|X\,S}P_{Z|Y})$, its secrecy capacity when CSI-T is non-causally available is denoted by $C_{10,s}(K_{SW})$. Finally, logarithms are taken to the base 2.

This research was supported in part by the European Commission under Marie Sklodowska-Curie Individual Fellowship No. 659316 (CYBERNETS) and the french Ministry of Defense through the Direction Générale de l'Armement (DGA).

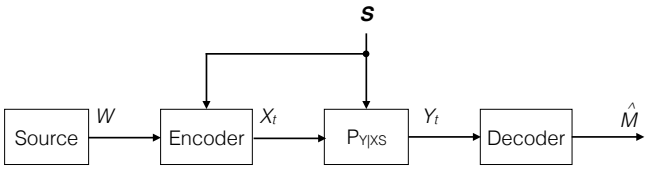

Fig. 1. State-dependent channel with non causal CSI-T, at channel use t.

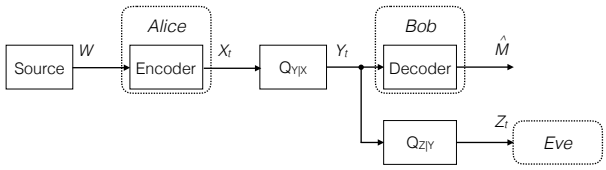

Fig. 2. Wiretap channel at channel use t.

2. STATE-DEPENDENT CHANNEL

A discrete memoryless state-dependent channel, denoted by $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$, describes a point-to-point communication in which the channel output Y depends on the channel input X and a state variable S. More specifically, at a given channel use t, the channel state $s_t \in \mathcal{S}$ is a realization of the random variable S. Hence, given the channel input $x_t \in \mathcal{X}$ and a state s_t, the channel output distribution is governed by the transition probability $P_{Y|X=x_t,S=s_t}$, see Fig.1.

Given a data transmission rate $R \geqslant 0$ and a block-length of n channel uses, the transmitter aims to send a message index $W \in \mathcal{W}$, with $\mathcal{W} = \{1, 2, \ldots, 2^{nR}\}$. This work exclusively focuses on the case of non-causal CSI-T, i.e., at the beginning of the transmission, the realization of the n-length state-sequence $\boldsymbol{S} = (S_1, S_2, \ldots, S_n)$ is known at the transmitter. The receiver is fully ignorant of the vector $\boldsymbol{S}$.

The encoder is defined by a deterministic function $f^{(n)} : \mathcal{W} \times \mathcal{S}^n \to \mathcal{X}^n$, which maps each pair $(W, \boldsymbol{S})$ into an n-length codeword $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$, i.e.

$$\boldsymbol{X} = f^{(n)}\left(W, \boldsymbol{S}\right). \tag{1}$$

The decoder is described by the function $\phi^{(n)} : \mathcal{Y}^n \to \mathcal{W}$ which, at the end of the transmission, outputs an estimate $\hat{W}$ of W using the received channel output vector $\boldsymbol{Y} \in \mathcal{Y}^n$. That is,

$$\hat{W} = \phi^{(n)}\left(\boldsymbol{Y}\right). \tag{2}$$

The decoding error probability, denoted by $P_{\mathrm{e}}^{(n)}$, is

$$P_{\mathrm{e}}^{(n)} = \Pr\left[\phi^{(n)}(\boldsymbol{Y}) \neq W\right]. \tag{3}$$

In the following the sets $\mathcal{S}, \mathcal{X}$, and $\mathcal{Y}$ are assumed to be finite. Within this context, an achievable rate is defined as follows.

Definition 1 (Achievable Rate for an SDC) *A rate R is said to be achievable for an SDC with non-causal CSI-T if for any $\epsilon > 0$ and for sufficiently large n, there exists a set of message indices $\mathcal{W} = \{1, 2, ..., 2^{nR}\}$, an encoding function $f^{(n)}$, and a decoding function $\phi^{(n)}$ such that $P_e^{(n)} \leq \epsilon$.*

Lemma 1 characterizes the capacity of the SDC with non-causal CSI-T, obtained by Gelfand and Pinsker in [6].

Lemma 1 (Capacity of an SDC [6, Theorem 1]) *Given an SDC denoted by $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$, the following holds*

$$C_{10}(K_S) = \max_{P_{UX|S}} \left[I(U;Y) - I(U;S)\right], \tag{4}$$

where U is an auxiliary random variable satisfying $|\mathcal{U}| \leq \min\left(|\mathcal{X}||\mathcal{S}|, |\mathcal{Y}| + |\mathcal{S}| - 1\right)$ and $P_{USX}(u, s, x)$ factorizes as $P_{USX}(u, s, x) = P_S(s) P_{UX|S}(u, x|s)$.

Given an input distribution $P_{UX|S}$, the set of achievable rates is denoted by $\mathcal{R}_{SDC}(P_{UX|S})$. $\bar{\mathcal{R}}_{SDC}$ denotes the union of $\mathcal{R}_{SDC}(P_{UX|S})$ over all input distributions in $\triangle(\mathcal{U} \times \mathcal{X})$.

3. WIRETAP CHANNEL

Consider the discrete memoryless WTC in Fig. 2. This channel is fully described by the tuple $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y\,Z|X})$. At channel use t, given an input x_t, the outputs Y_t and Z_t are governed by the joint distribution $Q_{Y\,Z|X=x_t}$. The output Y_t is observed by the legitimate receiver whereas the output Z_t is observed by the malicious receiver. In the following, the input and output alphabets $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ are assumed to be finite. For the ease of presentation, this analysis is restricted to the case of physically degraded WTCs, i.e., the transition probability of the channel factorizes as $Q_{Y\,Z|X} = Q_{Y|X} Q_{Z|Y}$.

Given a transmission rate $R \geq 0$ and a block-length n, the aim of the transmitter is to reliably send a message index $W \in \mathcal{W} = \{1, 2, \ldots, 2^{n\,R}\}$ to the legitimate receiver while keeping it secret with respect to the malicious receiver. The secrecy is measured in terms of the normalized leakage at the eavesdropper, i.e., $\frac{1}{n} I(W; Z_1, Z_2, \ldots, Z_n)$.

The encoder is described by a function $g^{(n)} : \mathcal{W} \times \mathcal{V}^n \to \mathcal{X}^n$, where the set $\mathcal{V}^n$ contains random n-length vectors following the distribution $Q_{\boldsymbol{V}}(\boldsymbol{V}) = \prod_{t=1}^{n} Q_V(v_t)$. Hence, given the message index W and the random vector $\boldsymbol{V}$, the corresponding codeword $\boldsymbol{X}$ is generated according to

$$\boldsymbol{X} = g^{(n)}(W, \boldsymbol{V}). \tag{5}$$

The decoder is described by a function $\psi^{(n)} : \mathcal{Y}^n \to \mathcal{W}$ that maps each channel output $\boldsymbol{Y}$ into an estimate of the message

$$\hat{W} = \psi^{(n)}\left(\boldsymbol{Y}\right). \tag{6}$$

The probability of error is defined as

$$P_{\mathrm{e}}^{(n)} = \Pr\left[\psi^{(n)}(\boldsymbol{Y}) \neq W\right], \tag{7}$$

and a secrecy rate is achievable if it complies with the following definition:

Definition 2 (Achievable Secrecy Rate for a WTC) *A rate R is said to be achievable for the WTC if for any $\epsilon > 0, \delta > 0$ and for sufficiently large n, there exists a set of message indices $\mathcal{W} = \{1, 2, ..., 2^{nR}\}$, an encoding function $g^{(n)}$, and a decoding function $\psi^{(n)}$ such that $P_e^{(n)} \leq \epsilon$, and $\frac{1}{n} I(W; \boldsymbol{Z}) \leq \delta$.*

The maximum secrecy rate is referred to as the secrecy capacity. The following lemma fully characterizes the secrecy capacity of a degraded WTC of the form $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X} Q_{Z|Y})$.

Lemma 2 (Secrecy Capacity of a WTC [1]) *Given a WTC denoted by $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X}Q_{Z|Y})$, the following holds:*

$$C_s(K_W) = \max_{Q_{VX}} \left[I(X;Y) - I(X;Z) \right]. \tag{8}$$

Given an input distribution Q_{VX}, the set of achievable rates is denoted by $\mathcal{R}_{WTC}(Q_{VX})$. $\bar{\mathcal{R}}_{WTC}$ denotes the union of $\mathcal{R}_{WTC}(Q_{VX})$ over all input distributions in $\triangle(\mathcal{V} \times \mathcal{X})$.

4. MAIN RESULTS

The main results of this paper establishes a duality between WTCs and SDCs. More specifically, it is shown that there are pairs of WTC and SDC with non-causal CSI-T for which the secrecy capacity of the former equals the capacity of the latter. Moreover, a secrecy capacity achieving scheme for the SDC is also a capacity achieving scheme for the WTC and *vice versa*.

As pointed out in [6], the capacity of the SDC $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$ with non-causal CSI-T decreases as the auxiliary random variable U in (4) and the state variable S become more correlated. Interestingly, a similar phenomenon occurs in the wiretap channel $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X}Q_{Z|Y})$. More specifically, when the channel output Z becomes more correlated with the channel input X, the secrecy capacity decreases. In the SDC, a binning technique is used to guarantee reliable communication, given that the CSI sequence $\boldsymbol{S}$ is unknown at the receiver. On the other hand, in the WTC, a binning technique is used to ensure both reliable and secret communication, given that full CSI is available. Under the condition that the amount of information shared between U and S in the SDC equals the amount of information shared between X and Z in the WTC, the same binning scheme can be used for both channels. Theorem 1 establishes that under certain conditions, a capacity achieving scheme for the SDC K_S achieves a fraction of the secrecy capacity of the WTC K_W.

Theorem 1 *Let $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$ be any SDC. Fix any input distribution P_X satisfying*

$$P_{UX|S}(ux|s) = P_{U|S}(u|s)\mathbb{1}_{\{x=\theta(u,s)\}}, \tag{9}$$

where U is an auxiliary random variable such that

$$|\mathcal{U}| \leq \min\left(|\mathcal{X}||\mathcal{S}|, |\mathcal{Y}| + |\mathcal{S}| - 1\right), \tag{10}$$

and $\theta : \mathcal{U} \times \mathcal{S} \to \mathcal{X}$ is a deterministic bijective mapping. Then, for any WTC $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X}Q_{Z|Y})$ with input distribution Q_{VX} satisfying

$$I(X;Z) = I(U;S), \tag{11a}$$

$$\text{and} \quad I(X;Y) \geq I(U;Y), \tag{11b}$$

it holds that

$$\mathcal{R}_{SDC}(P_{UX|S}) \subseteq \mathcal{R}_{WTC}(Q_{VX}). \tag{12}$$

Proof: The proof is detailed in section 5.1. ■

Theorem 2 establishes that under certain conditions, a secrecy capacity achieving scheme for the WTC K_W achieves a fraction of the capacity of the SDC K_S.

Theorem 2 *Let $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X}Q_{Z|Y})$ be any WTC. Fix a discrete finite set $\mathcal{V}$ and any input distribution Q_{VX}. Then, for any SDC $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$ with input distribution P_X satisfying* (9)-(11a)*, and*

$$I(U;Y) \geq I(X;Y), \tag{13}$$

it holds that

$$\mathcal{R}_{WTC}(Q_{VX}) \subseteq \mathcal{R}_{SDC}(P_{UX|S}). \tag{14}$$

Proof: The proof follows the same steps as that of Theorem 1, except that the first and second parts are switched. ■

Theorem 3 establishes that under certain conditions, the capacity of the SDC K_S coincides with the secrecy capacity of the WTC K_W.

Theorem 3 *Let $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X}Q_{Z|Y})$ be a WTC, and $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$ be an SDC satisfying* (9)-(11a)*, and*

$$I(U;Y) = I(X;Y), \tag{15}$$

for the capacity achieving input distribution $P_{UX|S}$ and the secrecy capacity achieving input distribution Q_{VX}. Then,

$$\bar{\mathcal{R}}_{WTC} = \bar{\mathcal{R}}_{SDC}. \tag{16}$$

Proof: Theorem 3 is a direct consequence of Theorem 1 and Theorem 2. Since (9)-(11a) and (15) are verified for the (secrecy) capacity achieving distributions, (12) and (14) yield (16). Note that it also follows that

$$C_{10}(K_S) = C_s(K_W). \tag{17}$$

■

The pairs of SDCs and WTCs for which Theorem 3 holds are said to be dual. For such a pair, the capacity achieving scheme for the SDC and the secrecy capacity achieving scheme for the WTC are the same. It follows that the encoders of the SDC and the WTC are exchangeable and guarantee the same achievable rates for both channels. Theorem 3 finds a straightforward application in determining the secrecy capacity of a special class of state-dependent wiretap channels.

Proposition 1 *Let $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$ be any SDC, and $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y|X}Q_{Z|Y})$ be a WTC satisfying* (9) - (11a) *and* (15) *for their (secrecy) capacity achieving input distributions. Then, the secrecy capacity of the state-dependent WTC $K_{SW} = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}, P_S, P_{Y|X\,S}Q_{Z|Y})$ with non-causal CSI-T is*

$$C_s(K_W) \overset{(a)}{=} C_{10}(K_S) \overset{(b)}{=} C_{10,s}(K_{SW}). \tag{18}$$

Note that this result crucially depends on the fact that the eavesdropper is ignorant of the CSI.

5. PROOFS

5.1. Proof of Theorem 1

The proof unfolds as follows. In the first part, a capacity achieving scheme for the SDC with non-causal CSI-T is described. In the second part, the same scheme is proved to achieve a fraction of the secrecy capacity of the WTC.

First part: Consider the SDC $K_S = (\mathcal{S}, \mathcal{X}, \mathcal{Y}, P_S, P_{Y|X\,S})$ and the following coding scheme.

Codebook generation: Consider the random variables $(U, S, X) \in \mathcal{U} \times \mathcal{S} \times \mathcal{X}$ with a fixed joint distribution $P_{USX}(u, s, x) = P_S(s)P_{U|S}(u|s)P_{X|US}(x|u, s)$, with $\mathcal{U}$ satisfying $|\mathcal{U}| \leq \min(|\mathcal{X}||\mathcal{S}|, |\mathcal{Y}| + |\mathcal{S}| - 1)$ and

$$P_{X|US}(x|s, u) = \mathbb{1}_{\{x=\theta(u,s)\}}, \tag{19}$$

where

$$\theta : \mathcal{U} \times \mathcal{S} \to \mathcal{X} \tag{20}$$

is a fixed deterministic bijective function. Fix a non-negative pair $(\bar{R}, R) \in \mathbb{R}^2$, with $\bar{R} \geq R$. For a given CSI vector $\boldsymbol{s} = (s_1, s_2, \ldots, s_n) \in \mathcal{S}^n$ and for each message $w \in \{1, 2, \ldots, 2^{nR}\}$, generate a set (sub-codebook) of $2^{n(\bar{R}-R)}$ i.i.d. vectors $\boldsymbol{u}(w, m) = (u_1(w, m), u_2(w, m), \ldots, u_n(w, m))$, with $m \in \{1, 2, \ldots, 2^{n(\bar{R}-R)}\}$, following the distribution $P_{\boldsymbol{U}|\boldsymbol{S}}$ defined as

$$P_{\boldsymbol{U}|\boldsymbol{S}}(\boldsymbol{u}|\boldsymbol{s}) = \prod_{i=1}^{n} P_{U|S}(u_i|s_i). \tag{21}$$

Encoding: To send the message index w knowing the CSI vector $\boldsymbol{s}$, the encoder chooses the sequence $\boldsymbol{u}(w, m)$ with the greatest m satisfying

$$(\boldsymbol{u}(w, m), \boldsymbol{s}) \in \mathcal{T}_\epsilon^{(n)}(U, S). \tag{22}$$

If such a jointly typical sequence does not exist, the encoder chooses $m = 1$. Note that this selection rule yields a binning structure in the codebook. At each channel use t, the encoder uses the function θ defined in (20) to generate the channel input $x_t = \theta(u_t(w, m), s_t)$, i.e.,

$$\boldsymbol{x}=(\theta(u_1(w, m), s_1), \theta(u_2(w, m), s_2), \ldots, \theta(u_n(w, m), s_n)).$$

The encoding function in (1) is therefore defined as

$$f^{(n)}(w, \boldsymbol{s}) = (\theta(u_1(w, m), s_1), \ldots, \theta(u_n(w, m), s_n)). \tag{23}$$

Decoding: At the end of the transmission, the receiver outputs the estimation $\hat{w}$ of w if there is a unique pair $(\hat{w}, \hat{m})$ satisfying $\left(\boldsymbol{u}(\hat{w}, \hat{m}), \boldsymbol{y}\right) \in \mathcal{T}_\epsilon^{(n)}(U, Y)$, with $\hat{m} \in \{1, 2, \ldots, 2^{n(\bar{R}-R)}\}$. Otherwise, the decoder outputs an error.

Probability of error analysis: An error might occur at the encoder or the decoder. Consider the following event observed at the encoder:

$$A_{w,m} : (\boldsymbol{u}(w, m), \boldsymbol{s}) \in \mathcal{T}_\epsilon^{(n)}(U, S). \tag{24}$$

Given the CSI vector $\boldsymbol{s}$ at the encoder, the event $A_{w,m}$ holds true if the indices w and m induce a codeword $\boldsymbol{u}(w, m)$ that is jointly typical with $\boldsymbol{s}$. Consider the following event observed at the decoder:

$$B_{w,m} : (\boldsymbol{u}(w, m), \boldsymbol{y}) \in \mathcal{T}_\epsilon^{(n)}(U, Y). \tag{25}$$

Given the channel output sequence $\boldsymbol{y}$ at the decoder, the event $B_{w,m}$ holds true if the indices w and m induce a codeword $\boldsymbol{u}(w, m)$ which is jointly typical with $\boldsymbol{y}$. Assume that the message index $w = 1$ is transmitted. Hence, by symmetry of the random coding argument, it holds that

$$\Pr\left[\hat{W} \neq W\right] = \Pr\left[\hat{W} \neq W | W = 1\right]. \tag{26}$$

More specifically, an error occurs if one of the following events holds true:

(a) None of the indices m satisfies that $(\boldsymbol{u}(1, m), \boldsymbol{s}) \in \mathcal{T}_\epsilon^{(n)}(U, S)$ at the encoder;

(b) Given that there exists at least an index m that satisfies $(\boldsymbol{u}(1, m), \boldsymbol{s}) \in \mathcal{T}_\epsilon^{(n)}(U, S)$ at the encoder, the codeword $\boldsymbol{u}(1, m)$ does not satisfy $(\boldsymbol{u}(1, m), \boldsymbol{y}) \in \mathcal{T}_\epsilon^{(n)}(U, Y)$ at the decoder;

(c) Given that there exists at least an index m that satisfies $(\boldsymbol{u}(1, m), \boldsymbol{s}) \in \mathcal{T}_\epsilon^{(n)}(U, S)$ at the encoder, there exist several codewords $\boldsymbol{u}(w', m)$, with $w' \in \mathcal{W} \setminus \{1\}$, that satisfy $(\boldsymbol{u}(w', m), \boldsymbol{y}) \in \mathcal{T}_\epsilon^{(n)}(U, Y)$ at the decoder.

Hence, from Boole's inequality, the following holds

$$\begin{aligned}
&P_{\mathrm{e}}^{(n)} = \\
&P_{\boldsymbol{S}}(\boldsymbol{s}) \Pr\left[\bigcap_{m=1}^{2^{n(\bar{R}-R)}} A_{1,m}^c \cup \bigcup_{m=1}^{2^{n(\bar{R}-R)}} \left((A_{1,m} \cap B_{1,m}^c) \cup \bigcup_{w=2}^{2^{nR}} B_{w,m}\right)\right] \\
&\leq \prod_{m=1}^{2^{n(\bar{R}-R)}} \Pr[A_{1,m}^c] + \sum_{m=1}^{2^{n(\bar{R}-R)}} \left(\Pr[A_{1,m} \cap B_{1,m}^c] + \sum_{w=1}^{2^{nR}} \Pr[B_{w,m}]\right) \\
&\leq \left(1 - 2^{-n(I(U;S)+\delta)}\right)^{2^{n(\bar{R}-R)}} + \epsilon + 2^{n\bar{R}} 2^{-n(I(U;Y)+\delta)} \\
&\leq \exp(-2^{n(\bar{R}-R+I(U;S)+\delta)}) + \epsilon + 2^{n\bar{R}} 2^{-n(I(U;Y)+\delta)}.
\end{aligned} \tag{27}$$

It follows that the probability of error $P_{\mathrm{e}}^{(n)}$ can be made arbitrarily small if

$$\bar{R} - R \geq I(U; S), \tag{28a}$$

$$\text{and} \qquad \bar{R} \leq I(U; Y). \tag{28b}$$

Thus,

$$R \leq \bar{R} - I(U; S) \leq I(U; Y) - I(U; S). \tag{29}$$

This completes the description of the coding scheme for the SDC. Now, the same encoder will be plugged in the WTC for the second part of the proof.

Second part: Consider the WTC $K_W = (\mathcal{X}, \mathcal{Y}, \mathcal{Z}, Q_{Y\,Z|X} = Q_{Y|X}Q_{Z|Y})$. Note that this channel model is not state-dependent. Here, the CSI vector $\boldsymbol{S}$ is replaced by a local source of randomness represented by an n-length random vector $\boldsymbol{V} \in \mathcal{S}^n$, distributed according to $P_{\boldsymbol{S}}(\boldsymbol{v}) = \prod_{i=1}^{n} P_S(v_i)$. Hence, to transmit the message index $w \in \mathcal{W}$, the transmitter uses the encoding function in (23) such that, $\boldsymbol{x} = f^{(n)}(w, \boldsymbol{v})$. In the following, each vector $\boldsymbol{v}$ is identified by an index $m \in \{1, 2, \ldots, 2^{n(\bar{R}-R)}\}$, and thus the local random vector can be denoted by $\boldsymbol{v}(m)$ to emphasize the index associated to each vector $\boldsymbol{v}$.

The remainder of the proof shows that using the coding scheme previously described allows to satisfy both the reliability constraint and the security constraint in the WTC.

Decoding: At the end of the n channel uses the legitimate receiver estimates the index pair $(\hat{w}, \hat{m})$ based on the received channel output vector $\boldsymbol{y}$. The decoder outputs $(\hat{w}, \hat{m})$ if it is the unique pair satisfying

$$\left(f^{(n)}(\hat{w}, \boldsymbol{v}(\hat{m})), \boldsymbol{y}\right) \in \mathcal{T}_\epsilon^{(n)}(X, Y). \tag{30}$$

Probability of error analysis: The probability of error analysis focuses on an upper-bound of the probability of error, defined in (7). Consider a third (virtual) receiver that observes

the transmitted message w as well as the channel output Z. Given w, the virtual receiver estimates m. An estimate $\bar{m}$ is a feasible estimation of m at the virtual receiver if it is the unique pair satisfying

$$\left(f^{(n)}(w, \boldsymbol{v}(\bar{m})), \boldsymbol{z}\right) \in \mathcal{T}_\epsilon^{(n)}(X, Z). \tag{31}$$

Assume without loss of generality that the index pair $(w, m) = (1, 1)$ is transmitted. Note that the probability of error $P_{\mathrm{e}}^{(n)}$ is upper-bounded by

$$P_{\mathrm{e}}^{(n)'} = \Pr\left[(\hat{W}, \hat{M}) \neq (1, 1) \vee \bar{M} \neq 1 | W = 1, M = 1\right]. \tag{32}$$

For each $(w, m) \in \{1, 2, \ldots, 2^{nR}\} \times \{1, 2, \ldots, 2^{n(\bar{R}-R)}\}$, consider the events

$$C_m\text{: } \left(f^{(n)}(1, \boldsymbol{v}(m)), \boldsymbol{z}\right) \in \mathcal{T}_\epsilon^{(n)}(X, Z), \tag{33}$$

$$D_{w,m}\text{: } \left(f^{(n)}(w, \boldsymbol{v}(m)), \boldsymbol{y}\right) \in \mathcal{T}_\epsilon^{(n)}(X, Y). \tag{34}$$

An error might occur if one of the following event holds true:
(a) The codeword $\boldsymbol{x} = f^{(n)}(1, \boldsymbol{v}(1))$ does not satisfy $(f^{(n)}(1, \boldsymbol{v}(1)), \boldsymbol{z}) \in \mathcal{T}_\epsilon^{(n)}(X, Z)$ at the virtual decoder;
(b) There exists at least one $m' \in \{2, \ldots, 2^{n(\bar{R}-R)}\}$, yielding the codeword $\boldsymbol{x} = (f^{(n)}(1, \boldsymbol{v}(m'))$ that satisfies $(f^{(n)}(1, \boldsymbol{v}(m')), \boldsymbol{y}) \in \mathcal{T}_\epsilon^{(n)}(X, Z)$ at the virtual decoder;
(c) The codeword $\boldsymbol{x} = f^{(n)}(1, \boldsymbol{v}(1))$ does not satisfy $(f^{(n)}(1, \boldsymbol{v}(1)), \boldsymbol{y}) \in \mathcal{T}_\epsilon^{(n)}(X, Y)$ at the legitimate decoder;
(d) There exists at least one $w' \in \mathcal{W} \setminus \{1\}$ or one $m' \in \{2, 3, \ldots, 2^{n(\bar{R}-R)}\}$, yielding the codeword $\boldsymbol{x} = f^{(n)}(w', \boldsymbol{v}(m'))$ that satisfies $(f^{(n)}(w', \boldsymbol{v}(m')), \boldsymbol{y}) \in \mathcal{T}_\epsilon^{(n)}(X, Y)$ at the legitimate decoder.
Hence, from Boole's inequality, the following holds:

$$\begin{aligned} P_{\mathrm{e}}^{(n)'} &= \Pr\left[C_1^c \cup D_{1,1}^c \cup \bigcup_{m=2}^{2^{n(\bar{R}-R)}} \left(C_m \cup \bigcup_{w=2}^{2^{nR}} D_{w,m}\right)\right] \\ &\leq 2\epsilon + \sum_{m=1}^{2^{n(\bar{R}-R)}} \left(\Pr[C_m] + \sum_{w=1}^{2^{nR}} \Pr[D_{w,m}]\right) \\ &\leq 2\epsilon + 2^{n(\bar{R}-R-I(X;Z)+\delta)} + 2^{n(R-I(X;Y)+\delta)}. \end{aligned} \tag{35}$$

Therefore, $P_{\mathrm{e}}^{(n)'}$ can be made arbitrarily small if

$$\bar{R} - R \leq I(X; Z), \tag{36a}$$

$$\text{and,} \qquad \bar{R} \leq I(X; Y). \tag{36b}$$

Consider now the leakage $\frac{1}{n} I(W; \boldsymbol{Z})$ at the eavesdropper:

$$\begin{aligned} &\frac{1}{n} I(W; \boldsymbol{Z}) \\ &\leq \frac{1}{n} I(\boldsymbol{U}, \boldsymbol{V}; \boldsymbol{Z}) - \frac{1}{n} I(\boldsymbol{V}; \boldsymbol{Z}) \\ &= \frac{1}{n} I(\boldsymbol{X}; \boldsymbol{Z}) - \frac{1}{n} I(\boldsymbol{U}; \boldsymbol{V}) - \frac{1}{n} I(\boldsymbol{Z}; \boldsymbol{V}|\boldsymbol{U}) + \frac{1}{n} I(\boldsymbol{U}; \boldsymbol{V}|\boldsymbol{Z}) \\ &= \frac{1}{n} I(\boldsymbol{X}; \boldsymbol{Z}) - \frac{1}{n} I(\boldsymbol{U}; \boldsymbol{V}) + \frac{1}{n} I(\boldsymbol{U}; \boldsymbol{Z}) + \frac{1}{n} I(\boldsymbol{U}; \boldsymbol{V}|\boldsymbol{Z}) \\ &\quad - \frac{1}{n} I(\boldsymbol{U}, \boldsymbol{V}; \boldsymbol{Z}) \qquad (37a) \\ &= \frac{1}{n} I(\boldsymbol{X}; \boldsymbol{Z}) - \frac{1}{n}(\bar{R} - R). \qquad (37b) \end{aligned}$$

A sufficient condition to minimize the leakage is

$$I(X; Z) - (\bar{R} - R) \leq \delta, \qquad \delta > 0. \tag{38}$$

Hence, to ensure reliable and secret communications, (36a) and (38) must be simultaneously satisfied. For instance, let R and $\bar{R}$ satisfy

$$\bar{R} - R = I(X; Z). \tag{39}$$

Plugging (39) into (36b) yields

$$R \leq I(X; Y) - I(X; Z). \tag{40}$$

Finally, from (11), any rate R satisfying (29) also satisfies (40), and is therefore an achievable rate for both channels. This yields (12) and completes the proof.

5.2. Proof of Proposition 1

First, note that since K_S and K_W satisfy (9) - (11a) and (15) for the (secrecy) capacity achieving distributions, equality *(a)* in (18) follows by applying Theorem 3. Note also that K_{SW} consists in appending the eavesdropper's channel considered in the WTC K_W to the channel K_S. Therefore, it holds that

$$C_{10,s}(K_{SW}) \leq C_{10}(K_S). \tag{41}$$

Equality *(b)* is established in the remainder of the proof. Consider the coding scheme presented in the proof of Theorem 1. It has already been shown in Section 5.1 that this scheme guarantees an arbitrarily low probability of error for any rate $R \leq C_{10}(K_S)$. Thus, it remains to show that this scheme also guarantees an arbitrarily small leakage at the eavesdropper for a rate R up to $C_{10}(K_S)$. Note that

$$\begin{aligned} I(W; Z) &\leq I(U, S; Z) - I(S; Z) \\ &\overset{(a)}{=} I(X; Z) - I(U; S) \\ &\overset{(b)}{=} 0, \end{aligned} \tag{42}$$

where *(a)* follows the same steps as (37a), and *(b)* is due to (11a). Thus, because of the assumption (11a), it follows that the leakage is guaranteed to be arbitrarily small for any achievable rate R. Therefore, this scheme guarantees both an arbitrarily small probability of error and an arbitrarily small leakage for rate R up to $C_{10}(K_S)$, which yields equality *(b)* in (18). This completes the proof.

6. CONCLUSION

A duality between WTCs and SDCs has been established in this paper. It has been shown that under certain conditions, achievable rates in the state-dependent channels with non-causal channel state information at the transmitter are achievable for wiretap channels with the same coding scheme and vice-versa. The proof is based on typicality arguments to show that the code construction described in [6] allows to satisfy the reliability constraint and the secrecy constraint in some wiretap channels. Moreover, it follows from this result that appending an eavesdropper satisfying the duality condition to a state-dependent channel has no impact on the (secrecy) capacity.

7. REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[3] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[4] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, no. 4, pp. 289–293, Oct. 1958.

[5] B. S. Tsybakov and A. V. Kuznetsov, "Coding in a memory with defective cells," *Problems of Information Theory*, vol. 10, no. 2, pp. 52–60, Apr. 1974.

[6] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, Jan. 1980.